ACADEMIC PRESS

J. Experimental Child Psychology 83 (2002) 53–75

Journal of Experimental Child Psychology

www.academicpress.com

Five-year-olds' handling of reference and description in the domains of language and mental representation☆

I.A. Apperly[a,*] and E.J. Robinson[b]

[a] *School of Psychology, University of Birmingham, Edgbaston, Birmingham, B15 2TT, UK*
[b] *Keele University, UK*

Received 2 October 2001; revised 24 July 2002

Abstract

Children's concurrent success on false belief tasks and their handling of two labels for one object (e.g., bunny/rabbit) has been interpreted as demonstrating understanding about the essential features of representation. Three experiments reveal the limitations in 5-year-olds' understanding for both mental and linguistic representations. We report relatively poor performance on a task involving two labels for one object (e.g., dice/eraser) which required children to treat another's knowledge as representing only some of the feature of its real referent: Dice but not eraser. Five year olds who made errors also had difficulty handling the fact that a written word 'dice' referring to such a dice/eraser, can also be applied to a standard dice but not to a standard eraser. These children lacked metalinguistic awareness of words as entities that both refer and describe.

Keywords: Metalinguistic awareness; Representations; Theory of mind

☆ The research reported was supported financially by the Economic and Social Research Council, UK Grant Number R000237932.
* Corresponding author.
E-mail address: i.a.apperly@bham.ac.uk (I.A. Apperly).

0022-0965/02/$ - see front matter
PII: S0022-0965(02)00102-9

Children's handling[1] of words as representations is linked to their performance on tasks that assess their handling of knowledge and beliefs as representations (e.g., Doherty & Perner, 1998; Doherty, 2000; Kamawar & Homer, 2000; Olson, 1994; Olson & Kamawar, 1999; Perner, 2000). This link sheds light on the nature of children's developing understanding of representations in different domains, and on the related practical problems of literacy and oral communication. Children's success on false belief tasks at around 4 years of age is widely held to mark a watershed in children's competence with mental representations, and consistent with this, research into their handling of linguistic and other external representations has mainly focused on developments at around this age. However, children's understanding of mental representations is in fact far from complete by 4–5 years, and so we might expect parallel later developments in their representational understanding of words. The identification of such developments in this paper takes us beyond previous findings concerning children's handling of mental and linguistic representations. In what follows, we begin by summarising the evidence for a link at around 4 years between handling of mental and linguistic representation, then we identify the limitations on children's abilities in the domain of knowledge and beliefs by showing what advances occur by around 6–7 years. From this we make predictions about related advances in metalinguistic ability, which are tested in the experiments we report.

By 4–5 years of age, children readily pass false belief tasks (e.g., Astington & Gopnik, 1988; Perner, Leekam, & Wimmer, 1987; Wimmer & Perner, 1983). Most three-year olds, in contrast, fail such tasks because they tend to predict a protagonist's action on the basis of their own knowledge of reality, or judge as if a protagonist's belief about some aspect of the world is in line with their own. Many authors treat children's developing abilities to reason about beliefs and knowledge as consisting in their coming to understand them as representations, and argue or assume that there is no further qualitative change in children's understanding in this domain (e.g., Flavell, 1988; Gopnik, 1993; Gopnik & Wellman, 1992; Perner, 1991, 2000 Perner et al., 1987).

At around the same time as children pass standard false belief tasks, they show similar advances in their handling of misrepresenting pictures (Robinson, Nye, & Thomas, 1994; Thomas, Jolley, Robinson, & Champion, 1999; Zaitchick, 1990), and in their metalinguistic awareness of some representational properties of language (Doherty & Perner, 1998). As a

[1] Where appropriate, we will talk in terms of children's "handling" of representations and representational properties, rather than children's "understanding." We do so because while older children and adults manifestly have practical mastery of the representational properties of words, pictures, and mental states, it is theoretically contentious whether such mastery derives from *understanding* representations as such.

measure of children's metalingustic awareness, Doherty and Perner (1998) assessed children's handling of synonyms (or rather terms that are assumed to be treated by children as synonymous). They engaged children in a game in which a puppet named a picture one way, and the child gave the other name. Three-year-olds performed poorly, while 4-year-olds performed well. Scores across a set of five such synonym production trials correlated strongly with children's performance on standard unexpected transfer false belief tasks. Doherty and Perner (1998) claim that these tasks are related because they make importantly similar demands on children's handling of representations. They argue that passing their synonyms task requires a distinction to be made between *what* is represented (word meaning) and *how* it is represented (in one particular linguistic form rather than another).

This distinction maps onto Perner's (1991) analysis of what is involved in acknowledging false belief. In a standard unexpected transfer task, for example, the child sees an object moved from one location to another in a protagonist's absence (e.g., Wimmer & Perner, 1983). To acknowledge false belief, Perner argues, children must differentiate what is represented (the object moved) from the particular way in which it is represented (as being in a particular location). Thus, according to Doherty and Perner, "both [synonym and false belief] tasks are based on a common conceptual understanding that one and the same state of affairs can be conceived of in different ways" (p. 303).

This general conception receives additional support from subsequent studies concerning children's understanding of homonymy (Doherty, 2000), and hypo- and hypernyms and alternative colour terms for the same object, though Perner (2000) questions whether all of these relations are, strictly speaking, tapping the same understanding of representations (see also Garnham, Brooks, Garnham, & Ostenfield, 2000 for an alternative interpretation). Other authors (Kamawar & Homer, 2000; Kamawar & Olson, 1999) follow a similar line on a common basis for children's metalinguistic awareness, and their developing understanding of mental representations.

However, success on false belief tasks does not indicate full competence with mental representations. Children cannot yet handle the fact that the very same input can be represented in different ways by different minds (Carpendale & Chandler, 1996; Chandler & Helm, 1984; Chandler & Sokol, 1999; Flavell, Speer, Green, & August, 1981; Robinson, 1994; Robinson & Robinson, 1982; Robinson & Whittaker, 1987; Sodian, 1988; Taylor, 1988). For example, 4- to 5-year-olds do not yet recognise that an ambiguous utterance can legitimately be interpreted in different ways by two different listeners (Carpendale & Chandler, 1996), that listeners might not know which interpretation the speaker intends (Robinson & Whittaker, 1987; Sodian, 1988) or that the viewer of a small uninformative part of an object might not know what it is (Robinson & Robinson, 1982; Taylor, 1988). A more precise identification of the circumstances under which children do

and do not over-estimate another's knowledge emerges from tasks developed by Apperly and Robinson (1998, 2001; Robinson & Apperly, 2001). We focus on these particular tasks because they lead to novel predictions about children's developing metalinguistic awareness that we shall test in the experiments reported below. Our argument will be that up to the age of 6 or 7 years children struggle to handle the fact that representations, whether mental or lingustic, both refer to their real referents and describe them in a particular way. Difficulty coordinating these two functions leads to related errors when children reflect on another's knowledge or about the meaning of words.

In Apperly and Robinson's (1998, 2001) tasks, the child participant knew an object under two descriptions, such as a story character who was a *neighbour* but also a *clown*, or a *ball* that was also a *present*. A protagonist was only partially informed, and knew the object under only one of the descriptions. Child participants were questioned in various ways about the protagonist's partial knowledge. We describe one task which involved a dice[2] that was also an eraser. A puppet, Heinz, saw the dice in a box but did not feel it so did not know it was also an eraser. This setting allowed investigation of children's handling of the fact that in descriptions of a person's knowledge or beliefs e.g., "Heinz knows that there's a dice in the box," the way in which the object of reference is described is commonly constrained by what that person knows, rather than what is really the case (see e.g., Quine, 1953). Hence, although in fact the dice is also an eraser, adults judge that it is wrong to substitute another description and say "Heinz knows that there's a eraser in the box," since Heinz does not know that the dice is an eraser (Apperly & Robinson, 1998). Descriptions that are constrained in this way are substitution-sensitive (referentially opaque). In contrast, the truth of many other types of sentence in which the dice/eraser might appear e.g., "It is true that there's a dice in the box," is unaffected by substituting another term to describe the object: It is true that there's an eraser in the box. In such sentences, these descriptions are substitution-insensitive (referentially transparent). Children's readiness to constrain the linguistic description of a person's knowledge in substitution-sensitive (referentially opaque) contexts offers an indirect test of their understanding that a person's knowledge only captures a subset of an object's actual features.

Using the dual identity materials described above, Apperly and Robinson (1998) asked children substitution-sensitive (referentially opaque) questions such as "Does Heinz know that there is a rubber in the box?" ("rubber" is British English for "eraser"). Questions of this kind posed significant problems for 5- and 6-year-olds, consistent with a much earlier study by Russell (1987), even after children were given contextual support to make the in-

[2] Although the correct singular term is "die" we follow 4–6-year-olds' convention of calling the object "a dice."

tended meaning clear (Apperly & Robinson, 2001). Importantly, children of this age did not seem to over-estimate knowledge when the test question was substitution-insensitive (referentially transparent): For example, ''Does Heinz know that the dice is a rubber?'' Children could answer such questions correctly, and found them no harder than standard false belief questions. From an adult perspective, children's responses to the two forms of question about Heinz's knowledge contradicted each other. The substitution-sensitive and substitution-insensitive questions apparently ask about the very same state of partial knowledge about the object of reference, yet the former are relatively hard and the latter relatively easy. Why?

One difference between the two types of question is that in the more difficult substitution-sensitive questions a single linguistic expression (''rubber'' in ''Does Heinz know that there's a rubber in the box?'') both *refers* to the object in question (the dice/eraser) and *describes* it in a particular way (as a rubber), that Heinz does not know. In contrast, in the relatively easy substitution-insensitive questions such as ''Does Heinz know that the dice is a rubber?'' one expression, ''dice,'' refers to the object (the dice/eraser), while another, ''rubber,'' describes it under the particular description that Heinz does not know.

Importantly the difficulties just identified are not just to do with children's use or understanding of language. Many children who make errors of over-estimation when judging descriptions of Heinz's knowledge of the dice-eraser, also predict wrongly that he will search for an eraser as if he was fully informed about the dice/eraser (Apperly & Robinson, 1998, 2001). Thus, children's difficulties seem to stem from the manner in which they represent for themselves Heinz's partial knowledge about the dice/eraser, and not just from ignorance or misunderstanding about the way adults use language. The suggestion is, then, that 4- and 5-year-old children have difficulty handling the fact that mental representations both refer to and describe their real referents in a particular way. We are lead to the expectation that children may also have parallel metalinguistic difficulties, which go beyond those described by Doherty and Perner (1998) and others.

Yet if, as Doherty and Perner (1998) claim, the link between children's handling of false beliefs and of synonyms at approximately 4 years resides in ''a common conceptual understanding that one and the same state of affairs can be conceived of in different ways,'' then 4- and 5-year-olds should already possess the ability to co-ordinate word reference and word description. ''Dice'' and ''eraser'' are just as surely different ways of viewing a state of affairs as ''chocolate'' and ''pencils'' (in the standard deceptive box task) and ''rabbit'' and ''bunny'' in Doherty and Perner's synonym tasks. Perhaps Apperly and Robinson's dual identity stimuli are intrinsically more difficult than Dohery and Perner's synonym stimuli. ''Dice'' and ''eraser'' are two contingently related aspects of the same object. This might pose more problems than labels, such as ''rabbit'' and ''bunny'' that identify the same

(or almost the same) aspect or meaning of a single object, or "chocolate" and "pencils" that identify alternative correct and incorrect identities for a single object. If this were the case, in a task analogous to Doherty and Perner's synonym production task, children should find it more difficult to produce an alternative label when it identifies a different aspect of an object (e.g., to say "dice" given "eraser") than when it is another term for the same aspect (e.g., to say "rabbit" given "bunny"). Alternatively, multiple descriptions for an object might pose no more problems for children than synonyms. That is, children who can generate alternative labels for the dual identity items like the dice/eraser, might still find it difficult to use the dual identity labels when they are in substitution-sensitive descriptions of someone's partial knowledge. This would leave open the possibility that the synonyms task does not exhaust children's metarepresentational understanding of words.

Experiment 1

We adapted Doherty and Perner's synonym production task (Doherty & Perner, 1998 Experiment 3) to include objects with a dual identity, such as a dice/eraser, among the original synonym stimuli. As in the original task, children heard a puppet give one name for an object, and had to produce the alternative label themselves. We also assessed children's handling of a character's partial knowledge of the same dual identity stimuli using tasks similar to those described above involving the dice/eraser (Experiment 1 in Apperly & Robinson, 1998). Would children perform equally well on the name production task with the dual identity stimuli, as with the original synonym materials? For the dual identity stimuli, how would children's performance in the name production task compare with their handling of the knowledge of a character who knew one name but not the other?

Method

Participants. Two children failed to complete the experiment because they refused to accept that a puppet could see, hear or feel anything. Twenty eight children entered the study. Sixteen were male and 12 female, aged 4:9–5:7 (mean 5:3). All children were from the same Junior/Infant school with a lower-middle class catchment area in Birmingham, UK, and spoke English as their first language.

Materials. In the *partial knowledge tasks* we used a puppet protagonist called Heinz, and items with dual identities (a dice/eraser and a ball/present), only one of which was visually apparent (dice and ball, respectively). Each was contained in a box. For the test phase of the *name production task*, there were two puppet protagonists: Moose and Tommy the tortoise. The dual

identity items were used again, and in addition four test pictures were presented on four separate sheets of paper. The pictures were of a woman/lady, a truck/lorry, a cup/mug and a television/TV. Eight further cartoon pictures were used in the pre-test vocabulary check.

Procedure. Children were tested individually in a room separate from the rest of the class. In the *partial knowledge tasks*, the procedure was similar to that used by Apperly and Robinson (1998, 2001). Children were first familiarised with the alternative identities for the dice/eraser and ball/present, and it was emphasised that only one was apparent from seeing. Next, Heinz the puppet was introduced, Heinz was then allowed to see a dual identity item in its box. Importantly, in the case of the dice/eraser he was not allowed to feel (so did not find out it was an eraser), and in the case of the ball/present, he was not told that it was a present. Children were then asked two questions about what Heinz knew: A substitution-sensitive question (e.g., "Does Heinz know there's a present in the box?") and a substitution-insensitive question (e.g., "Does Heinz know that the ball is a present?"). These questions were presented in counterbalanced order between-child.

The *name production task*, modelled closely upon Doherty and Perner's (1998) Experiment 3, began with a vocabulary check. Children were presented with four pictures on a piece of paper, and one dual identity item, placed on the same sheet of paper. Children were asked to point to the one named by the experimenter. All four pictures on each sheet were named, along with the dual identity item, but they were named in different spatial sequences across the four sheets to guard against children pointing to the correct picture by spatial anticipation. For this reason, the pictures and real item were named in a consistent order for each child. All four pictures and each dual identity item from the test phase appeared twice, allowing it to be named by two different labels on separate occasions.

Following the vocabulary check, children were introduced to the two puppets, Moose and Tommy, and told that they were going to play a naming game with them. "In this game, Moose will say one name for a picture and Tommy's job is to say another name; not the one Moose said but a different one. He's going to need some help from you." Moose was operated by the experimenter and "whispered" his name for the picture into the experimenter's ear. The experimenter relayed the name to the child: "Moose says [label X]; so what could Tommy say?" Tommy was given to the child to look after.

The game began with three warm up trials using items with obvious synonyms: A sofa/settee, a rabbit/bunny and a coat/jacket. The child was asked to "Look at this picture. Shall we see what Moose will call this one? He says 'sofa'. So what could Tommy say?" A few children were very quick to provide the correct answer. Most were not, and the experimenter followed with "Well, he could say 'settee' because settee is another name for sofa."

For the test items, children were told "Moose says [label X], so what could Tommy say?" If they did not respond they were asked, "What's

another name for [label X]?'' and finally ''Can you think of another way of saying [label X]?'' If a child repeated the name Moose had said they were told, ''That's what Moose said, Tommy's job is to say something different.'' If a child gave an inappropriate name, they were told that it was not a good name for the picture. In each case the above prompts were used to elicit another response. If these prompts failed to elicit a response, the experimenter said ''Well that's a tricky (meaning hard to British children) one isn't it, shall we try another?'' The six stimuli were always presented in the same order: Woman/lady; cup/mug; ball/present; dice/rubber; lorry/truck; TV/television. The particular label given by Moose was alternated independently for each stimulus, with the effect that both labels were given an equal number of times, but appeared with different combinations of other labels between children. The alternative label production task and partial knowledge tasks were run in counterbalanced order between-child.

Results

First, we compared children's performance on the substitution-sensitive and substitution-insensitive questions about the dice/eraser and ball/present to check that previous findings were replicated. Each child was given two scores out of 2 according to the number of correct judgments for each question type[3] (substitution-sensitive and substitution-insensitive), and an analysis of variance (ANOVA) was computed with question type as a repeated measure and age category (upper or lower half of the sample), question order (substitution-sensitive or substitution-insensitive first) and task order (knowledge judgement or name production task first) as between subject variables. Consistent with previous findings, children were more accurate on substitution-insensitive questions than on substitution-sensitive questions, $F(1, 20) = 25.9$, $p < .001$. All other effects were non-significant (all ps $> .436$).

Next, we considered children's performance on the *name production task*. There was some variability across the different stimuli with 27/28 (96%) children correctly producing the second name for the woman/lady, 20/28 (71%) for the cup/mug, 19/28 (68%) for the truck/lorry, 25/28 (89%) for the TV/television, 23/28 (82%) for the ball/present, and 27/28 (96%) for the dice/eraser. However, our main interest within the name production task was whether there was any difference in difficulty between producing synonyms and producing dual identity labels. For the dual identity stimuli, children were given a score of 0, 1, or 2 according to the number of alternative labels successfully produced (see mean in Table 1). With the synonym stimuli, children could produce a minimum of 0 and a maximum of 4 correct alternative

[3] As we have generally found in the past, there were no differences between children's accuracy of response to trials with the dice/eraser as compared to trials with the ball/present.

Table 1
Mean scores out of 2 ($N = 28$) for the four key test questions of experiment 1

Alternative name production task		Partial knowledge task	
Synonym stimuli	Dual identity stimuli	Substitution-insensitive question	Substitution-sensitive question
1.62 (.38)	1.79 (.42)	1.79 (.50)	1.00 (.90)

labels. To create a range comparable with that for the dual identity stimuli, these scores were divided by 2 (see Table 1). Scores on the dual identity and synonym stimuli were then entered as repeated measures into an analysis of variance (ANOVA) with age category and task order as between subject variables. There was no significant effect of label type, $F(1, 24) = 2.35$, $p = .138$. There was a main effect of age category $F(1, 24) = 19.9$, $p < .001$, with younger children performing less well than older children. There was also a significant interaction between task order and label type, $F(1, 24) = 7.62$, $p < .011$. t tests revealed that this was due to performance with the dual identity stimuli being superior when the name production task came first. The basis for this effect is unclear, but this has no implications for the conclusion that production of dual identity labels is no harder than production of synonym labels. There was no interaction between age category and label type ($p = .327$).

Finally, we were interested in whether there was any difference in difficulty between production of the dual identity labels in the name production task, and the substitution-sensitive questions of the partial knowledge task. An ANOVA was computed with scores on the name production task and substitution-sensitive questions as repeated measures, and age category and task order (name production or partial knowledge task first) as between subject variables. There was a significant main effect of task type $F(1, 24) = 17.1$, $p < .0001$, with better performance on name production. There was also a significant main effect of task order $F(1, 24) = 5.88$, $p = .023$, with performance on both tasks better when the name production task came first. All other effects were non-significant (all ps $> .079$).

Discussion

Our results with the synonym task are consistent with those of Doherty and Perner (1998; Perner, 2000) in that 4- and 5-year-olds found it relatively easy to respond to a puppet naming a picture by producing an alternative synonym. Importantly it was no more difficult for children to produce alternative labels for the dual identity stimuli than for synonym stimuli. Yet, near ceiling performance on this task contrasted with significantly lower performance on the substitution-sensitive question of the partial knowledge task. This suggests that children's difficulties with such substitution-sensitive

questions are not due to the particular characteristics of the dual identity stimuli. Rather, the partial knowledge task placed demands on children that are distinct from whatever is measured by the synonym tasks. The next question, then, is *how* the tasks differ in the demands they make on children's ability to handle "different ways" of conceiving the same object, and how this might be related to their developing understanding of linguistic representation. We begin by considering alternative accounts of Doherty and Perner's (1998) findings.

Doherty and Perner (1998) argue that their synonym task requires children to understand that the very same abstract meaning (e.g., the concept of RABBIT) may be represented in two different representational forms: "Rabbit" and "bunny." However, one might equally say that children are required to handle the fact that the very same concrete referent—a rabbit—may be identified with two different representational forms. Perner (2000) makes this point when he acknowledges that his analysis of synonyms cannot be applied to related problems that children have with hypo- and hypernyms, and different colour terms for the same object. Clearly, words such as "dog" and "animal" do not have the same meaning or sense, though they may share an object of reference. The same may be said for colour terms. If this alternative interpretation is correct, it leaves room for the possibility of further developments beyond 4 years of age in understanding about linguistic representation, which we shall now consider in more detail.

In the introduction, we summarised Apperly and Robinson's (1998, 2001) evidence concerning children's pattern of success and failure on substitution-insensitive and substitution-sensitive questions about another character's partial knowledge, along with their incorrect predictions about the partially informed character's search for a desired object. They concluded that by 5–6 years children do not just have a broad tendency to over-estimate another's knowledge. Rather, children have difficulty treating representations both as referring to an object and as describing it in a particular way. We predicted that there might be related and parallel difficulty with linguistic representations. Having argued above that the synonyms task may not be suitable for assessing this aspect of representational understanding, what we now require is a procedure that is suitable.

We designed a novel task, loosely based upon a moving word task developed by Bialystok (e.g., Bialystok, 1991, 2000). Children were first familiarised with a number of objects e.g., a green dice/eraser, a yellow dice/eraser a normal eraser, a normal dice and a pen. Next, a word (e.g., "dice") printed on a card, was introduced as "a word for" one of the dice/erasers. This set up reference to the dice/eraser object, under a particular description "dice." The word card was then moved among the other objects, and children's job was to judge whether it was legitimate for the word to go with each object in turn. The objects were chosen to stand in a variety of relationships with the referential and descriptive meanings of the word. In one case, it was possible

for both the descriptive and referential aspects of the word's meaning to be incompatible with the new object, as when the word "dice" began as a word for a dice/eraser and was moved to the pen. Alternatively, referential and descriptive aspects of the word's meaning could be fully compatible[4] with the new object, as when the same word moved to the second dice/eraser. Our main interest, though, was in children's judgements when the word moved to objects that bore a partial correspondence with the initial referent (e.g., dice and eraser, when the initial object was a dice/eraser). In such cases a clear match or mis-match with the descriptive meaning of the word ("dice") conflicted with the word's initial referential meaning (that it is a word for a dice/eraser). Thus, if children inappropriately confused the word's referential and descriptive meanings, they might sometimes deny that "dice" could go with a normal dice and/or accept that "dice" could go with a normal eraser. Examination of the relative difficulty of these four trial types should reveal a detailed insight into children's handling of the referential and descriptive functions of linguistic representations.

We also tested children on the same kind of partial knowledge task used in Experiment 1. Our interest was in whether the errors we expected to find on substitution-sensitive questions in the partial knowledge tasks would be related to any errors found on the metalinguistic task.

Experiment 2

Method

Participants. We tested 38 children, 21 boys and 17 girls, aged 5:0–6:0 (mean 5:5). All children were from a primary school with an upper-working class catchment area in Birmingham, UK, and spoke English as their first language.

Materials. The *partial knowledge tasks* used the same dice/eraser and ball/present materials, and Heinz the puppet as described in Experiment 1. The *moving word tasks* used the same dice/eraser and ball/present. Trials with the dice/eraser also used a differently coloured dice/eraser, a normal dice, a normal eraser and a pen. Trials with the ball/present also used a differently sized ball/present, an old looking ball, a gift wrapped "present" and a ruler. A warm-up trial used two toy cars, a toy bike and a toy aeroplane. All moving word trials involved a small plastic doll called "Jackie" who moved words written on small pieces of card. A reading pre-test was also conducted with similar word-cards.

[4] This is on the assumption that children do not treat the word as a proper name for the initial dice/eraser.

Procedure. Children were tested individually in a room separate from the rest of the class. The *partial knowledge tasks* followed the procedure described for Experiment 1.

The *moving word task* began with a reading pre-test examining children's ability to read the moving word stimuli: Dice, rubber, ball, present, bike, aeroplane, and car. It was emphasised to children that this was a difficult task, that it did not matter if they could not read the words, but that they should try their best anyway.

In the warm-up moving word task, children were introduced to a word game in which a puppet character (Jackie) found a word for one of a set of objects laid out on the table in front of them, before moving it to other objects. With each movement, the child was asked to say whether Jackie could put the word in that place. The child was told that we were going to find a word for the car (which the experimenter indicated by pointing rather than naming). A card with a word printed on it was introduced. Children were told that it said "car" and the word was placed on top of the car that the experimenter had indicated. Next, they were asked what the word said. After the child's response, the experimenter said "Here comes Jackie, where else could she put it?" The word was moved to the bike, and children were asked "Can it go here?" The word was then moved to a second car, followed by an aeroplane, before moving back to the first car. At each object, the child was asked "Can it go here?" or "What about here?" On the rare occasions when it was necessary, corrective feedback was given, for example by telling children that the word could not go with the aeroplane because the word said "car."

Test trials took a very similar form to the warm-up. Objects were introduced and laid out on the table in front of the child. Introduction to the two dice/erasers and ball/presents emphasised their dual identity while for the normal dice, normal eraser, normal ball and normal present children were told "This is just a dice," etc. The precise wording used by the experimenter varied somewhat between trials in this experiment (but see Experiment 3), and what follows is the typical wording. At the beginning of a typical trial with the dice/eraser the child was told "We're going to find a word for one of these [indicating items on table]. This word says 'dice' and it can go here [on a dice/eraser] because that's a word for this isn't it?" Children were then asked what the word said. Next, children were told that Jackie was going to move the word and see where else she could put it, and the word was moved among the other items in an order that was semi-counterbalanced between-child so that the word never moved directly from one dual identity item to another. At each item the child was asked "Can it go here?" or "What about here?" until the word was returned to its original location. The experimenter then suggested that they find a different word "Because this [dice/eraser] has got another name hasn't it?" Finding the second word, the experimenter said "Here's one, it says 'rubber' and that's another name (or 'word') for this isn't it?" and the movements

among the items was repeated. The whole procedure was then repeated for the ball/present. Order of presentation of the stimulus types and the label types was counterbalanced between-child. Order of presentation of the block of moving word tasks and the block of partial knowledge tasks was also counterbalanced between-child.

Results and discussion

First, we considered children's performance on the *moving word task*. In the reading pre-test 24/38 children could read none of the words, one child read 6 out of 7 words correctly, three children read 3 words, four read 2 words and six children read 1 word. Hence the general level of reading was very low, and these scores were not entered as factors in any analyses. Children's performance on the warm-up trial was very good, with almost all children accepting that the word "car" could move from one car to another, but not to an aeroplane or a bicycle. This suggested that children easily understood that the purpose of the game was to judge the appropriateness of the moving word for each of the objects with which it was placed.

On the test trials, a word (e.g., "dice") moved from a dual identity item (e.g. dice/eraser) to four other items and children were asked "Can it go here?" For the *totally non-matching* item (a pen or a ruler) the correct answer was always NO, for the *dual identity match* item the correct answer was always YES. There were two single identity items (e.g., a dice and an eraser) one with the *same label* as the word on the card (correct answer YES) and one with the *other label* of the dual identity item (correct answer NO). There were two dual identity stimuli, each with two labels, and thus a total of four questions of each type. Children were given a score of 0,1,2,3, or 4 according to their total number of correct answers to each question type.

We began by checking that children knew what the word on the card said when the card was initially placed on the first dual identity item, and that children would allow it to return to the original item at the end of each trial. Of 152 checks at the beginning of a trial, 145 were answered correctly, and when an incorrect answer was given, children were corrected and the task continued. When children were asked what the word said at the end of the trial performance was 152/152, suggesting that they had no difficulty remembering this essential premise. All data from all children were included in subsequent analyses

Next, we considered children's responses as the word moved between the four different items. As can be seen in Table 2, children's performance with the totally non-matching and dual identity match stimuli was high. This clearly suggested that children were able to follow the experimental task, and were prepared both to confirm and deny that the word could go in some locations. However, performance was lower with the same label and other label stimuli. An ANOVA was computed with scores on the different items

Table 2
Illustrative stimuli and performance in different conditions within the moving word task in experiments 2 and 3

Stimulus type	Example	Mean correct number of responses out of 4 (*SD*s)	
		Experiment 2	Experiment 3
Total non-matching	Pen	3.84 (.55)	3.95 (.22)
Dual identity match	Dice/eraser	3.87 (.41)	3.66 (.82)
Same label	Dice	3.55 (.83)	2.78 (1.33)
Other label	Eraser	3.05 (1.21)	3.05 (1.14)

Note. Examples are based upon the moving word being "dice" and the initial object being a dice/eraser.

(dual identity match, totally non-matching, same label and other label) as repeated measures, age category (upper or lower half of the sample) and task order (moving word or partial knowledge task first) as between subject factors. There was a significant main effect of item type, $F(1, 34) = 19.5$, $p < .0001$, with all other effects non-significant (all ps $> .22$). Post-hoc Bonferoni comparisons were made between the different item types. Children's performance was significantly worse when the word was moved to the other label item than when it was moved to the dual identity match or totally non-matching items ($p < .0001$ and $p = .003$, respectively). Although from Table 2 there appears to be a tendency for children to perform less well when the word was moved to the same label item, all other comparisons were non-significant (all ps $> .131$).

For the *partial knowledge task*, children were each given two scores of 0, 1, or 2 according to the number of correct answers given to the substitution-sensitive and substitution-insensitive questions. An ANOVA was computed with scores on these questions as repeated measures, age category (upper or lower half of the sample) and task order (moving word or partial knowledge first) as between subject variables. As in Experiment 1, substitution-insensitive questions were significantly easier than substitution-sensitive questions, $F(1, 34) = 17.1$, $p < .0001$, with mean scores of 1.53 (.80) and 1.00 (.93), respectively. There was also a significant effect of age (upper or lower half of the sample) $F(1, 34) = 7.2$, $p = .011$, with older children performing better than younger children, but there was no effect of task order or interactions of any kind. Again, the results are consistent with previous findings using this task (Apperly, 1999; Apperly & Robinson, 1998, 2001).

We were also interested in whether children's difficulties on the partial knowledge task were related to their performance on our modified moving word task. According to our analysis, the same label and other label conditions of the modified moving word task require children to co-ordinate the referential and descriptive functions of a single word. On the partial knowledge tasks, the substitution-sensitive question makes this demand, while the substitution-insensitive question does not. Our interest was therefore in whether children's performance on the substitution-sensitive question of

the partial knowledge task was related to their performance on the same label and other label conditions of the moving word task. For these purposes, children were given a single score, summing their performance on same label and other label trials of the moving word task, resulting in a maximum possible score of 8. There was a substantial correlation between this combined score and children's performance on the substitution-sensitive questions of the partial knowledge task ($r = .59$, $p < .001$) which remained significant when the effects of age were controlled ($r = .52$, $p < .001$). However, the strongest test of our hypothesis was that this correlation should be independent of children's performance on the substitution-insensitive question about partial knowledge, which did not require the child to co-ordinate the referential and descriptive functions of a single word. When performance on this question was controlled for, in addition to age, signs of a relationship remained, but it was only marginally significant ($r = .33$, $p = .051$).

In summary, the 5-year-olds in our modified moving word task responded correctly when both the referential and descriptive functions of the word were in mis-match with the new object (as when "dice" or "rubber" moved to the pen). Most children also responded correctly when both the referential and descriptive functions of the word matched the new object (as when "dice" or "rubber" moved to a second dice/eraser). This result suggested that they were not using a simple associative strategy of mapping words one-to-one with objects. These successes also made it clear that children were engaged, and understood the essential features of the game. Importantly though, the understanding underlying these successes was insufficient to allow a similar level of success when the referential and descriptive functions of the word were in conflict (as when "dice" moved to a dice that was not an eraser, or to an eraser that was not a dice). This clearly suggests that children's metalinguistic awareness of the representational properties of words is still developing at 5 years of age. However, given the relatively low overall incidence of errors, and the fact that the instructions issued by the experimenter varied somewhat between trials, it was important to replicate the findings with the modified moving word task before drawing firm conclusions.

The data also provide some support for the hypothesis that this metalinguistic development should be related to children's ability to answer questions about partial knowledge. The possible nature of this relation will be explored in more detail in the final discussion. However, in Experiment 3 we opted to focus exclusively upon our novel metalinguistic task and did not administer the partial knowledge task.

Experiment 3

In Experiment 3, we tested younger 4- and 5-year-olds in the hope of finding more errors than we did in Experiment 2. We also improved on

the procedure of Experiment 2 where there had been some ambiguity about whether a word was being selected as a name for a particular object, or whether an object was being selected to suit a particular word. In Experiment 3, we consistently and unambiguously identified the referent first, then introduced the word as a word for that particular referent.

Participants. Children ($N = 41$) aged between 4:6 and 5:5 ($M = 5{:}0$) were tested. All children attended a primary school with a lower-middle class/upper-working class catchment area in Birmingham UK, and spoke English as their first language.

Materials. These were the same as for the modified moving word task in Experiment 2.

Procedure. As before, children began with a reading pre-test, this time using only five words (car, dice, rubber, ball, and present). This was followed by a warm-up moving word task with corrective feedback when necessary. In the subsequent test trials, the objects were introduced and laid out on the table in front of the child. Introduction to the two dice/erasers and ball/presents emphasised their dual identity while for the normal dice, normal eraser, normal ball and normal present children were told "This is just a dice," etc. At the beginning of a trial with the dice/eraser the experimenter pointed to the dice/eraser and said "Now we're going to find a word for this; here's one." Having placed the word on the dice/eraser the experimenter said "It says 'dice' (or 'rubber' as appropriate) so it can go here because this rubber is a dice." Children were then asked what the word said. Next, Jackie moved the word among the other items in an order that was semi-counterbalanced between-child, avoiding moves directly from one dual identity item to another. At each item the child was asked "Can it go here?" or "What about here?" No feedback was given on these trials. Finally the word returned to the original dice/eraser, and after being asked "Can it go here?" the child was asked "What does this word say?" This process was repeated for the other label, "rubber." The whole procedure was then repeated for the ball/present. Order of presentation of the label types was counterbalanced between-children.

Results

In the reading pre-test, one child out of 41 was able to read 2 out of 5 words correctly, nine children were able to read 1 word, and 31 children were unable to read any words.[5] Since there was so little variation in these scores, they were not entered as a factor in the analysis. Performance on the warm-up task was very good: 35 out of 41 children accepted that the word

[5] Following the main experiment, the first 20 children were also tested on the reading sub-test of the British Ability Scales. However, only four of these children read enough words to reach baseline, and the test was abandoned.

"car" could move from one car to another, but not to an aeroplane or a bicycle, 5 out of 41 children made one error, and 1 out of 41 made 2 errors.

As in Experiment 2, children were given a score of 0, 1, 2, 3, or 4 according to their total number of correct answers with each stimulus type: Totally non-matching; dual identity match; same label; other label. These scores are detailed in Table 2.

We began by checking that children knew what the word on the card said when the card was initially placed on the first dual identity item, and when it finally returned at the end of each trial. There were 159 out of 164 correct answers to this question at the beginning of a trial, and when children made errors they were corrected and the task continued. There were 157 of 164 correct answers when the word was returned to the first dual identity item. Since inspection of the data revealed no systematic pattern in the commission of errors, these data were retained.

Next we considered children's responses as the word moved between the four different items. An ANOVA was computed with scores on the item types (dual identity match, totally non-matching, same label and other label) as repeated measures, and age category (upper or lower half of the sample) as a between subject factor. There was a significant main effect of item type, $F(1, 39) = 45.2$, $p < .0001$, with all other effects non-significant (all ps $> .087$). Post-hoc comparisons were made between the different item types, with a Bonferoni correction applied on the basis of six comparisons. Children made significantly more errors when the word was moved to the same label item than when it was moved to the dual identity match or totally non-matching items (both ps $< .0001$). Children also made more errors when the word was moved to the other label item than when it was moved to the totally non-matching item. Inspection of the means suggested a trend for other label trials to be harder than dual identity match trials, but this was non-significant once the Bonferoni correction had been made ($p = .116$). All other comparisons were non-significant (all ps $> .233$).

Following these gross comparisons, we considered the patterns of responses across the four key item types to distinguish the possible reasons for which children might be making errors on any one item. Out of 164 sets of responses to the four label conditions, in 14 cases (9%) children incorrectly judged that the word could go with the totally non-matching item. In 10 of these cases, children also made an error on the same label item, and in one case a child made errors on both the same label and totally non-matching items. In two other cases, children made errors on both the totally non-matching item and the other label item. These patterns were infrequent, and it is unclear whether they resulted from a positive strategy, or were unsystematic. Most errors (40% of total responses) consisted only of mistakes with the same label item, the other label item or both. These errors are the most interesting, since they are made by children who showed some metalinguistic awareness of word meaning in their correct responses to the

totally non-matching and dual identity match items. For these latter patterns, the frequency with which children made errors on the same label and other label stimuli was highly similar: 29 out of 164 and 27 out of 164 respectively, with 10 out of 164 children making errors on both. This pattern should caution against over-interpreting the differences between error rates to same label and other label items in the ANOVAs above and in Experiment 2, which include errors from children who are also making mistakes on totally non-matching and dual identity match items. Overall, the results of Experiment 3 support the conclusion from Experiment 2 that there are important shortcomings in 4- and 5-year-olds' handling of written words.

General discussion and conclusions

In the introduction, we described evidence from children's developing understanding of knowledge, beliefs, and language that suggests that many 4- to 5-year-olds have achieved some aptitude with representations. We also noted that evidence of later developments in children's understanding of knowledge and beliefs revealed the limitations of 4- to 5-year-olds' handling of mental representations, and raised the expectation that similar limitations would be found in their metalinguistic abilities.

In Experiment 1, we confirmed that 5-year-olds performed near ceiling in a game involving production of different names for one object, whether the names were synonyms such as rabbit/bunny or different labels for dual identity objects, such as dice/rubber. Doherty and Perner (1998) interpreted success on this task as revealing understanding of the distinction between what is represented and the particular way in which it is represented. Yet in the domain of mental representations, the particular conditions under which 5-year-olds tend to over-estimate a protagonist's knowledge of a dual identity object suggest that they still have difficulty representing for themselves the particular way in which the dice/eraser is represented by the protagonist.

In Experiments 2 and 3, we found parallel difficulties with written linguistic representations. Five-year-old children correctly rejected placement of a word by a new object when both the referential and descriptive functions of the word were inappropriate (movement to a totally non-matching item). Conditions in which both referential and descriptive functions of the word were appropriate for the new object (movement to a dual identity match item) were also relatively easy. However, children made significantly more errors on conditions in which there was some contradiction between the referential and descriptive functions of the words (movement to same label or other label items). This pattern clearly casts doubt upon the claim that 4- and 5-year-olds fully understand words *as representations* and that they possess metalinguistic awareness of the kind described by Doherty and Perner

(1998; see also Kamawar & Olson, 1999; Kamawar & Homer, 2000). Instead of tapping children's understanding of the representational relationship between words, objects and abstract meanings, perhaps Doherty and Perner's synonyms task requires some lesser insight into the many-to-one relationship between objects and words and among different words.

It has become widely recognised in recent years that conceptual development in 3- and 6-year olds should be considered in the context of the significant changes in working memory, executive, and inhibitory abilities over the same time period (e.g., Gordon & Olson, 1998; Hughes, 1998; Russell, 1996). Of particular relevance to the current findings is a study by Bialystok (1999) that reports a strong association between children's performance on her own moving word task and a measure of executive control (the dimensional change card sorting task, see e.g., Frye, Zelazo, & Palfai, 1995). This clearly raises the possibility that children's performance on such moving word tasks is primarily a measure of executive control rather than metalinguistic awareness. Like Bialystok's, our moving word task is likely to make executive as well as metalinguistic demands. Crucially though, executive and other basic performance demands are surely the same across our four conditions (dual identity match, totally non-matching, same-label, and other-label) and so cannot account for the fact that they differ in difficulty.

We now return to considering what 4- and 5-year-olds' metalinguistic abilities might amount to. It is interesting to consider in empirical terms what distinguishes mere linguistic aptitude from metalinguistic awareness of words as entities that both refer and describe. Imagine our moving word task without a concrete word to move. Instead, children would be asked to judge the validity of Jackie the doll's naming e.g., "dice" as she pointed at each stimulus in turn. In such circumstances, it would surely be surprising if children made the kind of errors found in Experiments 2 and 3. Although there was no such condition in our experiments, children did not appear to be confused by the repeated use of the same word when the experimenter introduced the varied set of stimuli, and indeed had any children failed to accept the validity of the labels, they would have been excluded from the data set. Brief reflection upon children's everyday language confirms that using the same words in different circumstances cannot present real difficulties. Thus, in their oral language, children appear already to be *using* words successfully to identify a particular meaning or aspect of an object. Yet, if children possess such ability, why do they make mistakes on our moving word task?

One possibility is that the presence of the concrete word token makes the mediating role of the word in the game explicit. There is considerable evidence from the literature on metalinguistic development that before the age of 5 or 6 children have little awareness of words as things in themselves. For example, Bialystok (2000) cites evidence that up to around the age of 6, children have difficulty judging the number of words in a sentence, or

to distinguish between long and short words (see also Gombert, 1992 for a comprehensive review). Thus, while children clearly use language and comprehend its use to perform acts of reference and description, they lack adult-like awareness of words as mediators of these activities. Against this background, we suggest that the introduction of a concrete, written word token adds to our moving word game *a word in itself*, making explicit the fact that it is this single, particular word that is moved among a variety of objects. The referential and descriptive functions that are implicit in the use of oral language are laid bare and this poses problems for the child who is now faced with two conflicting elements of meaning. On the one hand, with each use, the word has a referential relationship with a particular whole object (e.g., a dice/eraser); on the other, the word describes the object in a particular way (e.g., "dice") that determines which other objects it may label. We suggest that resolution of this conflict leads to a new metalinguistic awareness of words as entities that refer, but whose referential function is constrained by the particular terms of description.[6]

A further important question concerns the relation between these changes in children's understanding of words, and the existing literature on the development of 5- to 7-year-olds' understanding of representations in general. As mentioned in the introduction, children of this age have various difficulties understanding about intepretation of ambiguous input, and once they overcome these difficulties they have been characterised as having an "interpretive theory of mind" (e.g., Chandler, 1988; Carpendale & Chandler, 1996; Chandler & Sokol, 1999). We have argued elsewhere that it is unclear how such accounts might accommodate the discrepancies we have identified in children's ability to answer different kinds of question about partial knowledge (Apperly & Robinson, 1998, 2001) or about ambiguous utterances (Robinson & Apperly, 2001). If a child answers "Does Heinz know that there's a rubber in the box?" incorrectly, wrongly predicts that Heinz will search as if fully informed, yet answers "Does Heinz know that the dice is a rubber?" correctly, it is difficult to say whether or not she in possession of an interpretive theory of mind.

We believe that the necessary precision for describing children's difficulties will come from an account pitched at a lower level than changes in children's theories of the mind or of representations more generally. The current experiments investigated a candidate for this lower level difficulty. Our hypothesis that children might struggle with co-ordinating the referential and descriptive functions of words was based upon an important difference

[6] The notion of a word having a descriptive and referential function is clearly related to the philosophical notion of signs having senses and references. We use the less technical, intuitively based notion of referential and descriptive functions in the hope of avoiding the deep philosophical wrangles over the appropriateness and application of the sense-reference distinction.

between the substitution-insensitive questions about partial knowledge and ambiguous utterances that 5-year-old children can typically answer, and the substitution-sensitive questions that they find more difficult. Experiments 2 and 3 suggest that 4- and 5-year-old children do indeed have difficulty on a metalinguistic task where referential and descriptive word meanings are in conflict. This difficulty clearly does not reduce to these children lacking an interpretive theory of mind.

We believe that the current investigations have yielded a candidate explanation for the surprising discrepancies in children's ability to handle partial knowledge and ambiguous utterances. The findings of Experiment 2, where partial knowledge and modified moving word tasks were compared directly, were not entirely clear, and future studies will be necessary to examine this link more rigorously. More importantly, we hope that our analysis might provide a fruitful basis for future research on children's late developing understanding of mental and external representations, and the potentially important links with metalinguistic awareness and literacy (e.g., Olson, 1994).

References

Apperly, I. A. (1999). *Children's mental representation of referential relations: Representational partitioning and "theory of mind."* Unpublished Ph.D. Thesis, University of Birmingham, UK.

Apperly, I. A., & Robinson, E. J. (1998). Children's mental representation of referential relations. *Cognition, 63*, 287–309.

Apperly, I. A., & Robinson, E. J. (2001). Children's difficulties handling dual identity. *Journal of Experimental Child Psychology, 78*, 374–397.

Astington, J. W., & Gopnik, A. (1988). Knowing when you've changed your mind: Children's understanding of representational change. In J. W. Astington, P. L. Harris, & D. R. Olson (Eds.), *Developing theories of mind* (pp. 193–206). UK: Cambridge University Press.

Bialystok, E. (1991). Letters, sounds and symbols: Changes in children's understanding of written language. *Applied Psycholinguistics, 12*, 75–89.

Bialystok, E. (1999). Cognitive complexity and attentional control in the bilingual mind. *Child Development, 70*, 636–644.

Bialystok, E. (2000). Symbolic representation across domains in preschool children. *Journal of Experimental Child Psychology, 76*, 173–189.

Carpendale, J., & Chandler, M. (1996). On the distinction between false belief understanding and subscribing to an interpretive theory of the mind. *Child Development, 67*, 1686–1706.

Chandler, M. (1988). Doubt and developing theories of mind. In J. W. Astington, P. L. Harris, & D. R. Olson (Eds.), *Developing theories of mind* (1st ed., Vol. 1, pp. 387–414). Cambridge: Cambridge University Press.

Chandler, M., & Helm, D. (1984). Developmental changes in the contribution of shared experience to social role taking competence. *International Journal of Behavioural Development, 7*, 145–156.

Chandler, M. J., & Sokol, B. W. (1999). Representation once removed: Children's developing conceptions of representational life. In I. E. Sigel (Ed.), *Development of mental representation: Theories and applications* (pp. 201–230). Mahwah, NJ: Erlbaum.

Doherty, M. J. (2000). Children's understanding of homonymy: Metalinguistic awareness and false belief. *Journal of Child Language, 27*, 367–392.

Doherty, M., & Perner, J. (1998). Metalinguistic awareness and theory of mind: Just two words for the same thing? *Cognitive Development, 13*, 279–305.

Flavell, J. H. (1988). The development of children's knowledge about the mind: From cognitive connections to mental representations. In J. W. Astington, P. L. Harris, & D. R. Olson (Eds.), *Developing theories of mind* (pp. 244–270). UK: Cambridge University Press.

Flavell, J. H., Speer, J. R., Green, F. L., & August, D. L. (1981). The development of comprehension monitoring and knowledge about communication. *Monographs of the Society for Research in Child Development, 46*(5, Serial No. 192).

Frye, D., Zelazo, P. D., & Palfai, T. (1995). Theory of mind and rule-based reasoning. *Cognitive Development, 10*, 483–527.

Garnham, W. A., Brooks, J., Garnham, A., & Ostenfield, A. (2000). From synonyms to homonyms: Exploring the role of metarepresentation in language understanding. *Developmental Science, 3*, 428–441.

Gombert, J. E. (1992). *Metalinguistic development*. USA: University of Chicago Press.

Gopnik, A. (1993). How we know our minds: The illusion of first-person knowledge of intentionality. *Behavioural and Brain Sciences, 16*, 1–14.

Gopnik, A., & Wellman, H. (1992). Why the child's theory of mind is really a theory. *Mind and Language, 7*, 145–171.

Gordon, A. C., & Olson, D. R. (1998). The relation between acquisition of a theory of mind and the capacity to hold in mind. *Journal of Experimental Child Psychology, 68*, 70–83.

Hughes, C. (1998). Executive function in preschoolers: Links with theory of mind and verbal ability. *British Journal of Developmental Psychology, 16*, 233–253.

Kamawar, D., & Homer, B. D. (2000). Internal and external notions of metarepresentation: A developmental perspective. In J. W. Astington (Ed.), *Minds in the making: Essays in honour of David R. Olson* (pp. 197–211). Oxford: Blackwell.

Kamawar, D., & Olson, D. R. (1999). Children's representational theory of language: The problem of opaque contexts. *Cognitive Development, 14*, 531–548.

Olson, D. R. (1994). *The world on paper*. UK: Cambridge University Press.

Olson, D. R., & Kamawar, D. (1999). The theory of ascription. In P. D. Zelazo, J. W. Astington, & D. R. Olson (Eds.), *Developing theories of intention: Social understanding and self control* (pp. 153–167). Mahwah, NJ: Erlbaum.

Perner, J. (1991). *Understanding the representational mind* (1st ed). Cambridge, MA: MIT Press.

Perner, J. (2000). RUM, PUM, and the perspectival relativity of sortals. In J. W. Astington (Ed.), *Minds in the making: Essays in honour of David R. Olson* (pp. 212–232). Oxford: Blackwell.

Perner, J., Leekam, S. R., & Wimmer, H. (1987). Three year old's difficulty with false belief: The case for a conceptual deficit. *British Journal of Developmental Psychology, 5*, 125–137.

Quine, W. V. (1953). *From a logical point of view* (1st ed). Cambridge, MA: Harvard University Press.

Robinson, E. J. (1994). What people say, what they think and what is really the case: Children's understanding of utterances as sources of knowledge. In P. Mitchell & C. Lewis (Eds.), *Children's early understanding of mind: Origins and development* (1st ed., pp. 355–381). Hillsdale, NJ: Erlbaum.

Robinson, E. J., & Apperly, I. A. (2001). Children's difficulties with ambiguity and partial knowledge arise from problems with partial referential relations. *Cognitive Development, 16*, 595–615.

Robinson, E. J., Nye, R., & Thomas, G. V. (1994). Children's conceptions of the relationship between pictures and their referents. *Cognitive Development, 9*, 165–191.

Robinson, E. J., & Robinson, W. P. (1982). Knowing when you don't know enough: Children's judgements about ambiguous information. *Cognition, 12*, 267–280.

Robinson, E. J., & Whittaker, S. J. (1987). Children's conceptions of relations between messages, meanings and reality. *British Journal of Developmental Psychology, 5*, 81–90.

Russell, J. (1987). "Can we say...?" Children's understanding of intensionality. *Cognition, 25*, 289–308.

Russell, J. (1996). *Agency: Its role in mental development* (1st ed). Hove, UK: Erlbaum.

Searle, J. (1983). *Intentionality* (Vol. 1) (1st ed). Cambridge: Cambridge University Press.

Sodian, B. (1988). Children's attributions of knowledge to a listener in a referential communication task. *Child Development, 59*, 703–718.

Taylor, M. (1988). Conceptual perspective taking: Children's ability to distinguish what they know from what they see. *Child Development, 59*, 703–718.

Thomas, G. V., Jolley, R. P., Robinson, E. J., & Champion, H. (1999). Realist errors in children's responses to pictures and words as representations. *Journal of Experimental Child Psychology, 74*, 1–20.

Wimmer, H., & Perner, J. (1983). Beliefs about beliefs: Representation and constraining function of wrong beliefs in young children's understanding of deception. *Cognition, 13*, 103–128.

Zaitchick, D. (1990). When representation conflicts with reality: The preschooler's problem with false belief and "false" photographs. *Cognition, 35*, 41–68.